A Functional Form of The Zenga Curve Based on Rohde's Version of the Lorenz Curve

Muhammad Fajar[1,2*], Setiawan[2], Nur Iriawan[2] Eko Fajariyanto[1]

[1]Badan Pusat Statistik, Surabaya, Indonesia
[2]Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstrak. Kurva Zenga adalah alat untuk mengukur ketidakmerataan pendapatan yang merepresentasikan rasio pendapatan antara kelompok pendapatan bawah dan kelompok pendapatan atas. Kurva Zenga yang tepat adalah kurva Zenga yang dapat mendeteksi variasi pada rasio tersebut. Dalam paper ini, penulis menurunkan bentuk fungsional kurva Zenga yang berasal dari model kurva Lorenz versi Rohde. Hasil penelitian ini menyimpulkan bahwa bentuk fungsional kurva Zenga dari model kurva Lorenz versi Rohde adalah konstanta sehingga dia tidak dapat merepresentasikan fenomena ketidakmerataan yang sesungguhnya terjadi.

Abstract. The Zenga curve is a tool to measure income inequality that represents the income ratio between the bottom income group and the top income group. A proper Zenga curve is a Zenga curve that can detect variations in the Ratio. In this paper, we derive the functional form of the Zenga curve from Rohde's Lorenz curve model. The result of this paper is that the functional form of the Zenga curve from Rohde's version of the Lorenz curve model is a constant. It cannot represent the truly happening phenomenon of inequality.

1. Introduction

The Lorenz curve is a popular analytical tool used in research on income inequality. The Lorenz curve is a visual representation that can reflect income distribution inequality. This curve build by Lorenz [1] to represent the unequal distribution of wealth. Furthermore, the Lorenz curve can be expressed in the functional form (parametric model) that is not derived directly from inverse of the cumulative distribution function [2 - 7], but Gastwirth [8] formulated a Lorenz curve model formula that can be derived from inverse of the cumulative distribution function. The Gastwirth formula [8] has a weakness: the researcher must determine the statistical distribution fit for the data before fitting the Lorenz curve. On the other hand, several functional forms of Lorenz curves are known without first knowing the exact statistical distribution of the data, such as Lorenz Curve-Rohde [2], Lorenz Curve-Raasche [4], Lorenz Curve-Ortega [5], Lorenz Curve-Chotikapanich [9], etc. Rohde [2] suggested that his version of the Lorenz curve is more fit than the Lorenz curve model with one other parameter such as Lorenz Curve based Pareto Distribution [2], Lorenz Curve-Gupta [10], Lorenz Curve-Chotikapanich [9], Lorenz Curve-Kakwani [11].

In addition to the Lorenz curve, there is also a curve that represents the ratio of the average income of the lowest income group to the average income of the top income group called the Zenga curve. The Zenga curve can be derived from the Lorenz curve [12] [13]. The Zenga curve is more sensitive to detect changes in the income structure of the community than the Lorenz curve when there is an additional income effect on the community [14].

The authors argue that the Zenga curve may not detect the variation in ratio of the average income of the lowest income group to the average income of the top income group. It is presumably due to the Lorenz curve specification used in the income distribution modeling. Therefore, this study focuses on investigating the functional form of the Zenga curve of Rohde's Lorenz Curve. If the functional form of the Zenga curve is constant, the arguments previously described are proven. Rohde's Lorenz Curve was chosen based on the explanation at the end of the first paragraph.

2. Method

2.1. Rohde's version of the Lorenz Curve

Suppose $F(x) = P(X \leq x)$ represents the cumulative distribution function (CDF) of the random variable X, assuming the value of X is a positive (non-negative) value. Then $F^{-1}(p) = \inf\{x: F(x) \geq p\}$, it means that the inverse of CDF is a quantile. The Lorenz curve, $L(p)$ formulated as follows:

$$L(p) = q = \frac{1}{\mu_X} \int_0^p F^{-1}(s)\, ds$$

where μ_X is the mean of X. In this paper, X represents the expenditure/income of the population/households, q is the cumulative proportion of expenditure/income $(0 \leq q \leq 1)$ received by the cumulative proportion of the population or households $(p: 0 \leq p \leq 1)$.

The Lorenz curve must have the following characteristics:

$$\frac{dL}{dp} > 0, \qquad \frac{d^2L}{dp^2} > 0, \qquad L(p) = 0, \qquad L(0) = 0, \qquad L(1) = 1$$

Rohde [2] proposed the Lorenz curve model, $L_R(p; \beta)$, as follows:

$$L_R(p; \beta) = \frac{(\beta - 1)p}{\beta - p}, \beta > 1, 0 \leq p \leq 1. \tag{1}$$

Meanwhile, to estimate the parameter in equation (1), the author uses the method proposed by Castillo, et al. [15] based on the point $(p_i, q_i), i = 1,..,n$ on the empirical Lorenz curve:

$$\hat{\beta}_i = \frac{p_i(1 - q_i)}{p_i - q_i} \tag{2}$$

It happens that $p_i = q_i = 0$ that causing the value of $\hat{\beta}_i$ cannot be defined so that it must be deleted which implies that the number of $\hat{\beta}_i$ is reduced by one $(n - 1)$ in equations (4), (5), dan (6). Estimation of β [15] in equation (1) is:

$$\hat{\beta}_M = \frac{1}{n - 1} \sum_{i=1}^{n-1} \hat{\beta}_i \tag{3}$$

$$\hat{\beta}_{Med} = \text{Median} \left(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{n-1} \right) \tag{4}$$

$$\hat{\beta}_{LS} = \frac{\sum_{i=1}^{n-1} p_i(1 - q_i)(p_i - q_i)}{\sum_{i=1}^{n-1}(p_i - q_i)^2} \tag{5}$$

2.2. Zenga Curve

The Zenga curve is formulated as follows [11]:

$$Z(p) = \frac{p - L(p)}{p\big(1 - L(p)\big)}. \tag{6}$$

The Zenga curve $Z(p)$ measures the inequality between the bottom $100p\%$ of the population against the top $100(1- p)\%$ by comparing the mean expenditures of the two groups. Equation (6) shows that the Zenga curve can be derived from the Lorenz curve. Then based on equation (6), the Zenga index can also be formulated as follows:

$$\zeta = \int_0^1 Z(p)dp. \tag{7}$$

3. Results and Application

By combining equation (1) into equation (6), it will get:

$$Z_R(p) = \frac{p - L_R(p)}{p(1 - L_R(p))} = \frac{p - \dfrac{(\beta - 1)p}{\beta - p}}{p \left(1 - \dfrac{(\beta - 1)p}{\beta - p} \right)}$$

$$= \frac{\dfrac{p(\beta - p) - (\beta - 1)p}{\beta - p}}{p \left(\dfrac{\beta - p - (\beta - 1)p}{\beta - p} \right)} = \frac{p(\beta - p) - p(\beta - 1)}{p(\beta - p - (\beta - 1)p)}$$

$$= \frac{p\big((\beta - p) - (\beta - 1)\big)}{p(\beta - p - (\beta - 1)p)} = \frac{\big((\beta - p) - (\beta - 1)\big)}{(\beta - p - (\beta - 1)p)}$$

$$= \frac{-p + 1}{\beta - p - \beta p + p} = \frac{1 - p}{\beta - \beta p} = \frac{1 - p}{\beta(1 - p)} \tag{8}$$

So, equation (8) can be simplified to:

$$Z_R(p) = \frac{1}{\beta} \tag{9}$$

Equation (9) is a function in the form of a constant, meaning that regardless of the value of p, then $Z_R(p)$ always has a value of $1/\beta$. The functional form of equation (9) is also known

as the uniform function or constant. However, it should be noted that in equation (8) the condition $\beta \neq p$ must be fulfilled. Then based on equation (7), the Zenga index ζ_R (10) is obtained whose formulation is the same as equation (9). It's happening because equation (9) is a constant.

$$\zeta_R = \int_0^1 Z_R(p)dp = \int_0^1 \frac{1}{\beta} dp = \frac{1}{\beta} \int_0^1 dp = \frac{1}{\beta}(p|_0^1) = \frac{1}{\beta} \qquad (10)$$

The application of the Zenga curve formulation and the Zenga index using the research results of Ref. [16]:

$$L_R(p;\beta) = \frac{0.485p}{1.485 - p} \qquad (11)$$

Equation (11) uses $\hat{\beta}_{Med}$ in equation (4) because the estimation results produce a minimum mean squared error (MSE) compared to $\hat{\beta}_M$ (3) and $\hat{\beta}_{LS}$ (5) in the case of this study. Based on equation (10) the Zenga curve and Zenga index can be derived from equations (9) and (10) are:

$$\zeta_R = Z_R(p) = \frac{1}{1.485} = 0.673. \qquad (12)$$

The interpretation of the Zenga curve (12) derived from Rohde's version of the Lorenz curve is that the average household income at each level p of the lowest income group is 67.3% lower than the average income of all levels of the top income group of the population. The Zenga curve derived from Rohde's Lorenz curve has a weakness, namely that it assumes that the ratio of income between the lowest and the top group for all levels is constant (uniform). This does not reflect the reality of what happened. Supposedly, the ratio varies at each level of *p*. Because the Zenga curve depends on the Lorenz curve, the correct functional form of the Lorenz curve is the key so that the Zenga curve represents the phenomenon.

4. Conclusions

Based on the material, it said that the functional form of the Zenga curve from Rohde's Lorenz curve model is a constant with the condition that $\beta \neq p$, but it has not reflected the reality. It implies that the Zenga index has the same formulation as its functional model. The correct functional form of the Lorenz curve is the key so that the Zenga curve represents the real phenomenon inequality.

References

[1] M. O. Lorenz, "Methods of measuring the concentration of wealth," Publications of the American Statistical Association, vol. 9, no. 70, pp. 209-219, 1905.

[2] N. Rohde, "An alternative functional form for estimating the Lorenz curve," Economics Letters 105, pp. 61-63, 2009.

[3] N. Rohde, "Lorenz curve interpolation and the Gini coefficient," Journal of Income Distribution, vol. 19, No. 2, pp. 111 – 123, 2010.

[4] R. Rasche, J. Gaffney, A. Y. C. Koo, and N. Obst, "Functional forms for estimating the Lorenz curve," Econometrica 48, pp. 1061–1062, 1980.

[5] P. Ortega, G. Martin, A. Fernández, M. Ladoux, and A. García, "A new functional form for estimating Lorenz curves," Review of Income and Wealth, series 37, number 4, 1991.

[6] T. Ogwang and U.L.G. Rao. Chotikapanich, "A new functional form for approximating the Lorenz curve", Economics Letters, vol. 52, no. 1, pp. 21 – 29, 1996.

[7] S. Paul and S. Shankar, "An alternative single parameter functional form for Lorenz curve," Empirical Economics, vol. 59, pp. 1393 – 1402, 2020.

[8] J. L. Gastwirth and M. Glauberman, "The interpolation of the Lorenz curve and Gini index from grouped data," Econometrica, vol. 44, no. 3, pp. 479-483, 1976.

[9] D. Chotikapanich, "A comparison of alternative functional forms for the Lorenz curve," Economics Letters, vol. 41, pp. 129-138, 1993.

[10] M. R. Gupta, "Functional form for estimating the Lorenz Curve", Econometrica, vol. 52, issue. 5, pp. 1313-1314

[11] N. C. Kakwani and N. Podder, "On the estimation of Lorenz curves from grouped observations," International Economic Review, vol. 14, no. 2, pp. 278-292, 1973.

[12] M. Langel and Y. Tillé, "Inference by linearization for Zenga's new inequality index: a comparison with the Gini index, "Metrika International Journal for Theoretical and applied Statistics 75, no. 8, pp. 1093–1110, 2012.

[13] M. Zenga, "Inequality curve and inequality index based on the ratios between lower and upper arithmetic means," Statistica e Applicazioni 4:3–27, 2007.

[14] W. Maffenini and M. Polisicchio, "How potential is the I(p) inequality curve in the analysis of empirical distributions, Technical report," Universita Degli Studi di Milano-Bicocca, 2010.

[15] E. Castillo, A. S. Hadi, and J. M. Sarabia, "A method for estimating Lorenz curves," Communications in Statistics, Theory and Methods, vol. 27, pp. 2037-2063, 1998.

[16] M. Fajar, "Pemodelan Kurva Lorenz versi Rohde pada Pengeluaran Rumah Tangga Pertanian di Provinsi Papua," Euclid, vol. 8, no.1, pp. 1-5, 2021.